[BoldFont=Fira Sans SemiBold]Fira Sans Book Fira Mono

popquizbox[1] colback=nipun-lightblue!10, colframe=nipun-blue, boxrule=2pt, arc=3pt, left=8pt, right=8pt, top=8pt, bottom=8pt, title= **Quick Quiz 1**, fonttitle=**, coltitle=nipun-white, colbacktitle=nipun-blue, enhanced, attach boxed title to top left=xshift=0pt, yshift=-2pt, boxed title style=arc=3pt, boxrule=0pt**

definitionbox[1] colback=nipun-green!8, colframe=nipun-green, boxrule=1.5pt, arc=2pt, left=6pt, right=6pt, top=6pt, bottom=6pt, title= **Definition: 1**, fonttitle=**, coltitle=nipun-white, colbacktitle=nipun-green**

examplebox[1] colback=nipun-orange!8, colframe=nipun-orange, boxrule=1.5pt, arc=2pt, left=6pt, right=6pt, top=6pt, bottom=6pt, title= **Example: 1**, fonttitle=**, coltitle=nipun-white, colbacktitle=nipun-orange**

keypointsbox colback=nipun-blue!8, colframe=nipun-blue, boxrule=1.5pt, arc=2pt, left=6pt, right=6pt, top=6pt, bottom=6pt, title= **Key Points**, fonttitle=**, coltitle=nipun-white, colbacktitle=nipun-blue**

$$P(x_1, x_2, x_3, \ldots, x_N|y) = P(x_1|y)P(x_2|y)\ldots P(x_N|y)$$

$$P(x_1, x_2, x_3, \ldots, x_N|y) = P(x_1|y)P(x_2|y)\ldots P(x_N|y)$$

Why is Naive Bayes model called Naive?

$$P(x_1, x_2, x_3, \ldots, x_N|y) = P(x_1|y)P(x_2|y)\ldots P(x_N|y)$$

Why is Naive Bayes model called Naive?

Naive assumption x_i and x_{i+1} are independent given y

$$\text{i.e. } p\left(x_2 \mid x_1, y\right) = p\left(x_2 \mid y\right)$$

$$
\begin{aligned}
&P(y = 1|w_1 = 0, w_2 = 0, w_3 = 1) \\
&= \frac{P(w_1 = 0|y = 1)P(w_2 = 0|y = 1)P(w_3 = 1|y = 1)P(y = 1)}{P(w_1 = 0, w_2 = 0, w_3 = 1)} \\
&= \frac{0.6 \times 0.8 \times 0.6 \times 0.5}{Z}
\end{aligned}
$$

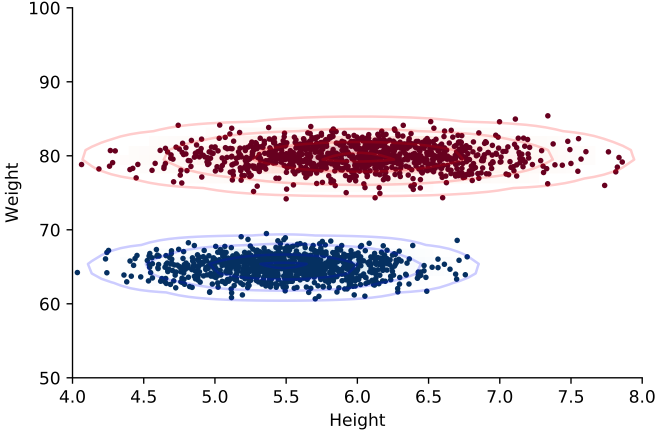
$$P(y = 1|w_1 = 0, w_2 = 0, w_3 = 1)$$
$$= \frac{P(w_1 = 0|y = 1)P(w_2 = 0|y = 1)P(w_3 = 1|y = 1)P(y = 1)}{P(w_1 = 0, w_2 = 0, w_3 = 1)}$$
$$= \frac{0.6 \times 0.8 \times 0.6 \times 0.5}{Z}$$

Similarly, we can calculate
$P(y = 0|w_1 = 0, w_2 = 0, w_3 = 1) = \frac{0.6*0.4*0.6*0.5}{Z}$

$$
\begin{aligned}
&P(y=1|w_1=0,w_2=0,w_3=1)\\
&= \frac{P(w_1=0|y=1)P(w_2=0|y=1)P(w_3=1|y=1)P(y=1)}{P(w_1=0,w_2=0,w_3=1)}\\
&= \frac{0.6 \times 0.8 \times 0.6 \times 0.5}{Z}
\end{aligned}
$$

Similarly, we can calculate
$P(y=0|w_1=0,w_2=0,w_3=1) = \frac{0.6*0.4*0.6*0.5}{Z}$
$\frac{P(y=1|w_1=0,w_2=0,w_3=1)}{P(y=0|w_1=0,w_2=0,w_3=1)} = 2 > 1$. Thus, classified as a spam example.

Note: no cross covariance! Remember all features are independent.



Wikipedia Example

Height	Weight	Footsize	Gender
6	180	12	M
5.92	190	11	M
5.58	170	12	M
5.92	165	10	M
5	100	6	F
5.5	100	6	F
5.42	130	7	F
5.75	150	7	F

Example

	Male	Female
Mean (height)	5.855	5.41
Variance (height)	3.5×10^{-2}	9.7×10^{-2}
Mean (weight)	176.25	132.5
Variance (weight)	1.22×10^{2}	5.5×10^{2}
Mean (Foot)	11.25	7.5
Variance (Foot)	9.7×10^{-1}	1.67

Classify the Person

- Given height = 6ft, weight = 130 lbs, feet = 8 units, classify if it's male or female.

Classify the Person

- Given height = 6ft, weight = 130 lbs, feet = 8 units, classify if it's male or female.
- $P(F|6ft, 130lbs, 8units) = \frac{P(6ft|F)P(130lbs|F)P(8units|F)P(F)}{P(130lbs, 8units, 6ft)}$

Classify the Person

- Given height = 6ft, weight = 130 lbs, feet = 8 units, classify if it's male or female.
- $P(F|6ft, 130lbs, 8units) = \dfrac{P(6ft|F)P(130lbs|F)P(8units|F)P(F)}{P(130lbs, 8units, 6ft)}$

Classify the Person

- Given height = 6ft, weight = 130 lbs, feet = 8 units, classify if it's male or female.
- $P(F|6ft, 130lbs, 8units) = \dfrac{P(6ft|F)P(130lbs|F)P(8units|F)P(F)}{P(130lbs, 8units, 6ft)}$
- $P(130lbs|F) = \frac{1}{\sqrt{2\pi\times550}} \times \exp\frac{-(132.5-130)^2}{2\times550} = .0167$

Classify the Person

- Given height = 6ft, weight = 130 lbs, feet = 8 units, classify if it's male or female.
- $P(F|6ft, 130lbs, 8units) = \dfrac{P(6ft|F)P(130lbs|F)P(8units|F)P(F)}{P(130lbs, 8units, 6ft)}$
- $P(130lbs|F) = \frac{1}{\sqrt{2\pi \times 550}} \times \exp \frac{-(132.5-130)^2}{2\times 550} = .0167$
- Finally, we get probability of female given data is greater than the probability of class being male given data.